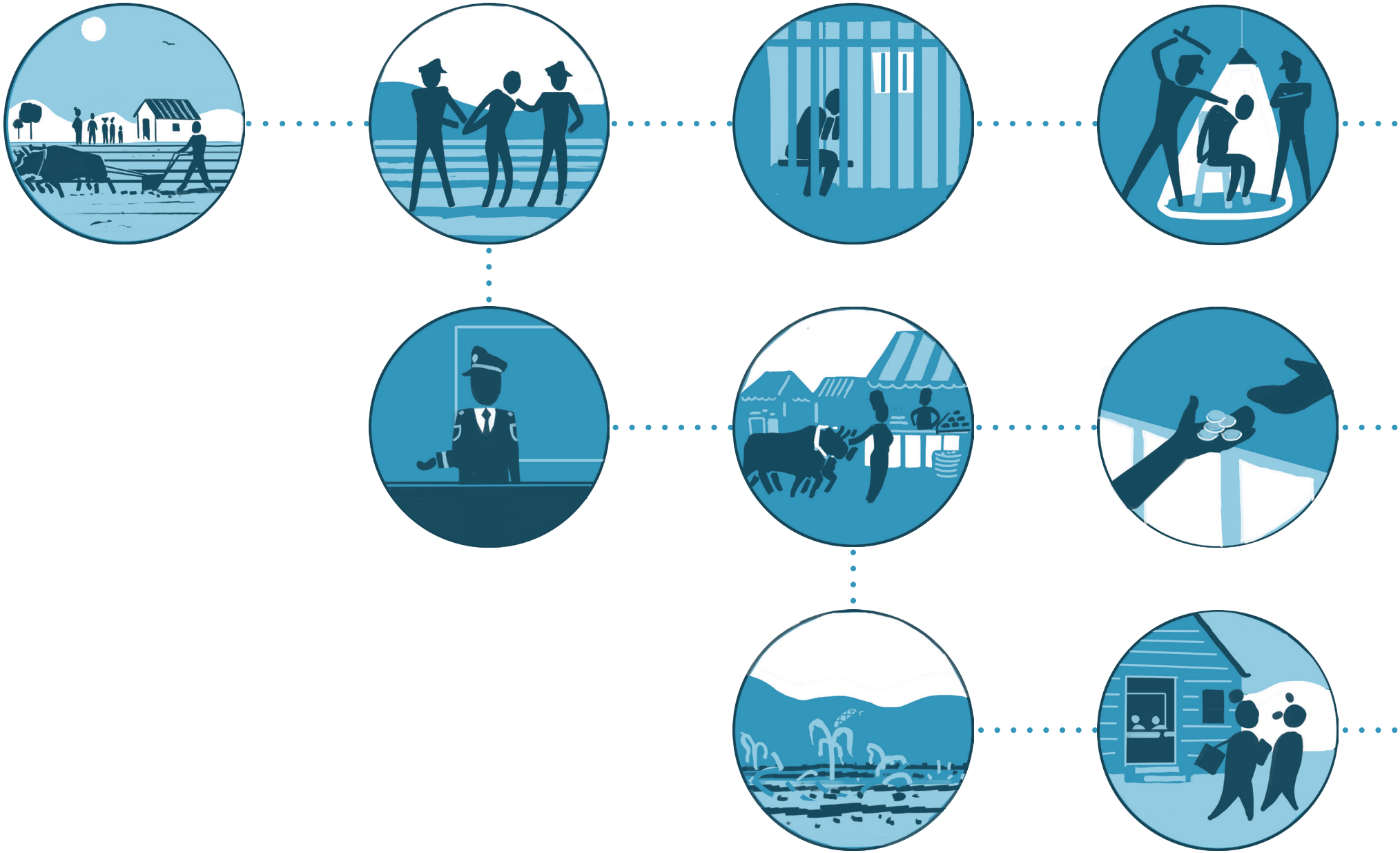
3 Data Viz Questions



How to Use This Booklet

Data visualization can make information more memorable, more persuasive, facilitate understanding and ultimately motivate action. And within human rights research, it can help investigators and researchers draw a bigger picture from individual human rights abuses by allowing them to identify patterns that may suggest the existence of abusive policies, unlawful orders, negligence, or other forms of culpable action or inaction by decision-makers.

The purpose of this activity is to explore some of the ideas and principles around designing effective data visualization for human rights advocacy.

This activity is broken into a series of six topics each with its own PDF guide.

In practice not every visualization process follows these exact steps in this same order. However, for a workshop setting, we present these as a way to walk through the topics. Each topic has a corresponding list of options and choices. Read through each topic and follow the instructions. Explore the options for each step as you progress.

The six steps are:

Step 1: Choose a human rights issue

Step 2: Discuss some kinds of data you might acquire

Step 3: Consider what question are you trying to answer with your data and visualization

Step 4: Choose a chart type for your visualization

Step 5: Consider some data and visualization hazards

Step 6: Consider some ways your charts can be improved

Step 3

What question are you trying to answer?

Considering your data, what are some questions you would like to illuminate with your data visualization? How does your data illustrate the respect for, protection of, or fulfillment of specific human rights? What elements of the rights you are examining are reflected in your data, and what is left out? Is your data better suited to function as structure, process, or outcome indicators? And what kinds of questions would best suit advocacy around your human rights issue?

Since human rights issues have many dimensions worth exploring, try to come up with long list of questions, particularly questions that do not have just one single answer.

For this step, it will always be important to have a clear sense of how your data might relate to the legal norms you are exploring. A number of frameworks have been developed to relate quantitative and qualitative data to human rights legal standards. Some of these are generic frameworks, potentially applicable to any right. Others are rights-specific, and still others are better suited to some kinds of rights than others. These rubrics might be helpful as you develop questions that you seek to answer through data analysis. Some examples include:

- The UN Office of the High Commissioner for Human Rights has published Human Rights Indicators: A Guide to Measurement and Implementation, which includes a standard format for constructing human rights indicators. This framework recommends that you first identify the various attributes of a given right, then develop indicators to measure structure, process, and outcome for each attribute. The different types of indicators will call for different kinds of data. This framework can be used for any human right.

- The Center for Economic and Social Rights (CESR) has developed the OPERA Framework for monitoring the fulfillment of economic, social, and cultural rights. This framework recommends that you identify quantitative indicators to measure Outcomes, or the level of realization of a given right; qualitative and quantitative indicators to assess Policy Efforts taken by a state to realize the right; budget and other data relevant to analyzing whether the state is devoting adequate Resources to the right; and methods for undertaking a contextual Assessment of any constraints on the state before drawing conclusions about the state's overall compliance with the relevant human rights standard. This framework was developed for

use with data relevant to economic, social, and cultural rights.

- The "Who Did What to Whom" framework was developed by Dr. Patrick Ball while he was at the American Association for the Advancement of Science. It is aimed at helping human rights workers design projects aimed at "the representation of acts of violence in a way that allows human rights researchers to make systematic, comparative analyses of patterns of violation in time, space, and social structure." While the model is set out as a how-to guide for researchers, it can also be used by those seeking to understand the possibilities and limits of a given events-based data set. Although it was developed for civil and political rights (especially violent acts), the model is relevant for any rights in which specific acts count as violations (e.g., extrajudicial executions, forced evictions, and unjustified water shut-offs).

3 | QUESTIONS

What are the patterns?

Is there regularity within the data? What are the frequencies? Where do things diverge from the regularity?

What is the distribution? How are amounts or qualities distributed across another attribute, like time?

What are the deviations? What data points do not follow the pattern?

What are the proportions? How do the parts compare to each other? Showing disproportionate impact may be particularly useful for establishing a human rights abuse or fulfillment issue.

What are the rankings? What is the best? What is the worst?

Sometimes patterns can be revealed by using a single encoding, like position along a single axis, though combining multiple encoding may reveal additional insights.

Patterns of human rights abuse may reveal a lack of protection or a specific effort to target a population, or the disproportionate impact of a practice, policy, or law on particular groups or areas.

3 | QUESTIONS

What are the correlations?

In seeking to determine a course of action or intervention that may yield a particular outcome, it may help to start by trying to identify correlations. In statistics, a correlation is a number that describes the degree of relationship between two variables.

A strong correlation means that a change in one data set closely tracks to a change in another data set. A weak correlation indicates that changes in the data sets are not very closely related. While correlation cannot, on its own, reveal causation, it can be useful in exploring possible variables for deeper causal analysis.

Scatter-plots are particularly effective for illustrating correlations. With one variable along the x-axis and another along the y-axis the existence and egree of correlation can be visualized, and the coherence of clusters and trends can be observed.

A challenge for human rights advocates is that correlation may be influenced by additional or "hidden" forces. See the section on Spurious Correlation in booklet five on visualization hazards and improvements.

3 | QUESTIONS

What is the change over time?

In some cases, raw values may be less important than relative changes over time.

What are the trends? Related to change over time, trends can also highlight velocity, how quickly something is changing.

What is the difference before and after? Not just change over time, but at two points in time—often before and after an event or a legal or policy change.

Socio-economic data can be compared over time to indicate progress or deterioration in the protection of rights and whether disparities are increasing or decreasing. Simple counts or event-based data collected at different points can also show trends over time.

How does it work?

What are the processes or elements of a system?

What are the components? What are the steps in the sequence? What are the inputs and outputs? What are the contexts and influences?

A flow chart, timeline, or pictorial illustration may work well to illuminate how things work. Given that a process unfolds over time, small multiple graphics may be useful to show changes from one state to the next.

Some examples include:

- illustrating judicial proceedings showing the stages in the path of a case
- diagramming how budgeting and procurement processes work
- depicting the impacts of particular policies and the chain of events resulting from this

What are the qualities?

Can the qualities be measured or categorized?

What are the similarities and differences? What is the comparison? Visual differences can reveal different inputs, outcomes or impacts.

What are repeated ideas, concepts, or conceptual categories?

One example would be disaggregating a body of data by demographic categories such as sex, race, income, or migration status to show disparate impact of government policies on different groups, or to explore disparate outcomes such as under-5 mortality or school dropout rates.

Using multiple visualizations to visualize different states can also help illuminate differences.

Note, once categories are created, they have a tendency to remain stable and can influence further analysis, as changing categories complicates the examination of change over time.

What are the amounts?

One basic question when working with quantities is simply, how much?

To what extent? What are the intensities? What are the ranges? What are the outliers?

Examining economic, social, and cultural rights, this could be the amount of an indicator. For example, examining fulfillment of the right to education, this could be the count of schools, of class sizes, of student enrollment, of students graduating, of functioning school furniture, etc.

Examining civil and political rights, this could be the number of incidents or victims: the number attacked, killed, or tortured, the number of prisoners, missing, or exiled.

Documenting very large-scale civil and political rights abuses can be challenging particularly across a large geographic area. Using appropriate methodologies are essential for accurate data collection, data processing, and analysis.

Data visualization provides a number of tools for exploring and interpreting quantitative data, such as setting the scales and ranges, “binning” or combining ranges of data into more managable groups, or adding interactive ways to filter data by attribute.

What are the relationships?

Networks are a special kind of system composed of entities that have relationships with each other. These are often visualized as network diagrams using nodes (representing entities) connected by lines representing relationships.

Are the relationships hierarchical? If so, what are the hierarchies? When establishing accountability, it may be important to track a chain-of-command, a supply chain, or a chain of beneficial ownership.

What are the relationships between the part and the whole? Where are the relationships between “parent” and “child” elements?

Tree diagrams, network diagrams, mind maps, and flow charts are some useful chart types for visualizing relationships. Visualizing and navigating large-scale networks of hundreds or thousands of nodes is a particular challenge.

3 | QUESTIONS

What are the locations?

Where did a human rights violation occur? What locations or regions are affected by a particular policy? What is the difference between urban and rural areas? How are things distributed spatially?

Displaying points on a map can contextualize where specific incidents took place. In addition to contextualizing specific events, mapping can reveal spatial patterns and gaps. Tools like choropleth maps, heat maps, and satellite images are useful for displaying geographic information.

Geographic information can also be a useful tool for finding correlations between two data sets, for instance funding of schools and the locations of different populations.

However, just because the data is geographic does not mean it needs to be displayed on a map. For instance, a ranked list or scatter plot may be more appropriate depending on your analysis and story.

3 | QUESTIONS

What is the movement or flow?

What are the mechanics of a given social or institutional system? How do resources flow with the commercial sector through government ministries to benefit citizens? How does information flow within an organization? How do resources or people flow between places? How do refugees or migrant workers flow between places? What are the obstacles along the way?

Sankey diagrams, chord diagrams, flow charts, flow maps, connection maps, and network diagrams are some tools use to depict flow.

After diagramming a set of movements or flow, additional layers can be added to show other kinds of interactions, for instance what kinds of laws affect different parts of the flow, what are different types of corruption found at points along the flow, or what parts of the flow are cut off or slowed.

3 | QUESTIONS

What are the structures in place to respect, protect, and fulfill a right?

Structural indicators measure the existence of institutions, legal rules, and other mechanisms needed to implement human rights.

What laws, constitutional provisions, or accountability mechanisms can you identify, count, or categorize?

How do local, state, and national structures compare with international and regional structures?

What formal, informal, and customary institutions exist that have a role in protecting rights?

Legal structures can be visualized across jurisdictions, perhaps coded, graded, or categorized by impact on a map, cartogram or chart.

Visualization can also help make legal structures legible and can help audiences understand their rights and the stages of legal procedures.

3 | QUESTIONS

What are the processes used to respect, protect, and fulfill a right?

Process indicators measure the efforts of governments and other duty bearers to implement human rights laws and make their enjoyment real.

What are the availability, accessibility, and quality of processes and interventions?

What institutions and processes exist to resolve disputes? What is the trust in courts? Are complaints about police misconduct satisfactorily resolved? What proportion of those accessing food aid are members of a national minority group?

Some of this is structural and may be visualized by mapping out work-flows and procedures. However, perception is also crucial to procedural fairness: what is the experience and opinion of individuals and groups seeking justice? Such qualitative information may be captured in a survey.

3 | QUESTIONS

What are the outcomes that can be used to measure the realization of a right?

Tracking outcome indicators over time provides a measure of the progressive realization of a right. (Though outcomes alone do not give a full picture of state compliance with human rights obligations.)

Development outcomes such as maternal mortality or primary school completion can be used to help assess levels of fulfillment of the human right to health or education.

Comparing disaggregated data over time can reveal disparities in the rates of progress among different groups of people.

Development indicators may be convenient but should be scrutinized as much as other types of data. Uncovering the methodology of data collection and processing may make some indicators problematic for human rights use. Where possible, affected communities should participate in the selection of indicators that meaningfully reflect their experience.

3 | QUESTIONS

What are the resources devoted to an issue?

How are resources allocated and spent? What are state or institutional budgets and expenditures? What are capital, material, or personnel expenses? How transparent is this?

What resources are available? How are resources generated? And from which segments of the population? What types of taxes are collected, and who avoids paying? What is the national debt or foreign aid?

Which populations benefit from resources? How have resources changed over time? How does resource generation and expenditure compare to other similar countries?

It is important here to also understand constraints of the state, including domestic or international limitations on resources. These policy or process limitations may have significant impacts, as well as the conduct of third parties such as international lending conditions, trade rules, donor restrictions on aid, or domestic corruption or mismanagement.

About this Booklet

This activity is the result of a research collaboration between the Center for Human Rights and Global Justice and Tandon School of Engineering at New York University, and was funded with a grant from the John D. and Catherine T. MacArthur Foundation.

The booklet was authored by John Emerson and Margaret Satterthwaite with help from contributors Brianne Cuffe and Sidra Mahfooz.

It was inspired by Shiqing He and Eytan Adar's Vizit cards, The Data Visualisation Catalogue by Severino Ribecca, and Tamara Munzner's Nested Model for Visualization Design and Validation.

For more information about data visualization and human rights along with links to resources, research and tools, visit our project page at http://visualizingrights.org.

Version 0.1, published October 2017.

Cover illustrations by John R. Holmes for the Open Society Justice Initiative report *Presumption of Guilt: The Global Overuse of Pretrial Detention* (2014.)

Please send suggestions, comments, or feedback to john@backspace.com.